Moral Intuitions and Their Reliability

Morgan Lawless

TC 660H

Plan II Honors Program

The University of Texas at Austin

May 15, 2019

Paul Woodruff, Ph.D.

Department of Philosophy

Supervising Professor

Hannah Wojciehowski, Ph.D.

Department of English

Second Reader

Abstract

Philosophers often defer to intuition when discussing ethics. It is common to evaluate an ethical theory by comparing its suggested action in a dilemma to what intuition suggests. A theory that aligns with intuition is said to be more credible, while one that does not align is said to be less credible. But in other cases, philosophers suggest discarding intuition in favor of what theory prescribes. The trouble with these practices is knowing when to trust intuition and when to trust theory. Philosophers and psychologists give convincing arguments to doubt the reliability of ethical intuitions, but I come to their defense in this paper. I argue that intuitions are reliable enough to use as starting points to form ethical beliefs. I explain how understanding the relationship between intuitions and beliefs can help settle moral disagreement. Lastly, I attempt to improve upon the standard resolution to the intuition versus theory problem, known as reflective equilibrium.

Contents

1 Introduction

What is the right thing to do? How should you act? And how are you supposed to know? A large part of the typical answers to these questions involve moral intuitions. Intuitions are the gut feelings we are all familiar with that pop up in all aspects of life, not just in ethics. In ethics, these might come in the form of gut feelings or unconsciously formed "seemings" by which someone just knows or feels that something is the right or wrong thing to do. We all use our moral intuitions to navigate the typical decisions of everyday life and judge more complex dilemmas. We appeal to others' intuitions when trying to persuade them that something is right or wrong. "Come on, if he did that to you, how would you feel? You should know better." Everyday appeals to intuitions such as this one are actually complicated ethical arguments that raise a series of questions. How much weight should we give to our moral intuitions? Are they reliable and track the moral "truth"? Or are moral intuitions themselves "true"? How do you settle disagreement between differing intuitions? Can there be such a thing as a moral expert with more reliable intuitions? These questions are at the root of moral philosophy and their answers have deep and extensive impacts on how we ought to live out every aspect of our lives.

Moral philosophers have generally favored ethical theories that expand upon intuitions. Ethical theories are considered strong if they align with moral intuitions at least somewhat, but most theories also ask people to give

up some of their intuitive judgements in favor of what the theory prescribes. The most well-known ethical theories are utilitarianism (a form of consequentialism), deontology (duty- or rule-based), and virtue ethics. Proponents of each of these theories point out how they often align with common moral intuitions. Detractors will critique these theories by bringing up examples where the theory recommends an action that is highly counter-intuitive, implicitly arguing that the intuition is correct. These common philosophical moves raise the question, under what circumstances should we defer to theories and under what others should we defer to intuitions? If intuitions always reign supreme, then why bother creating ethical theories? If theory is always the way to go, then why do philosophers so often reference intuition when evaluating a theory? The answer seems to be that both intuition and theory are important, but how to tell which prevails when is anything but obvious.

This paper explores these questions and others about intuitions. To do so, I start by trying to find a suitable definition for what I mean when I use the word "intuition". From there, I take a look at various proposals for whence intuitions come. After analyzing the nature of intuitions, I offer a variety of skeptical arguments that deny the reliability of intuitions, followed by some defenses of reliability. Ultimately, I will make the case that some intuitive reasoning can be reliable and be the foundation for moral beliefs. I argue that cases where intuitions conflict with theory can also be understood as moral beliefs in conflict. While no foolproof method exists to know which belief is correct, there are heuristics to guide the choice.

2 What are intuitions?

Identifying the nature of intuitions is a difficult task because regular people and academic philosophers refer to a variety of distinct concepts, all under the umbrella term "intuition". While this paper's main focus is on ethical intuitions, this section is applicable to intuitions generally.

Also keep in mind that I use the words "ethical" and "moral" interchangeably throughout this paper to mean the same thing. While there are colloquial differences between "ethics" and "morality", I use both to mean one thing: how one ought to act. A system of ethics/morality is instructions for how one ought to act.

2.1 Are intuitions beliefs?

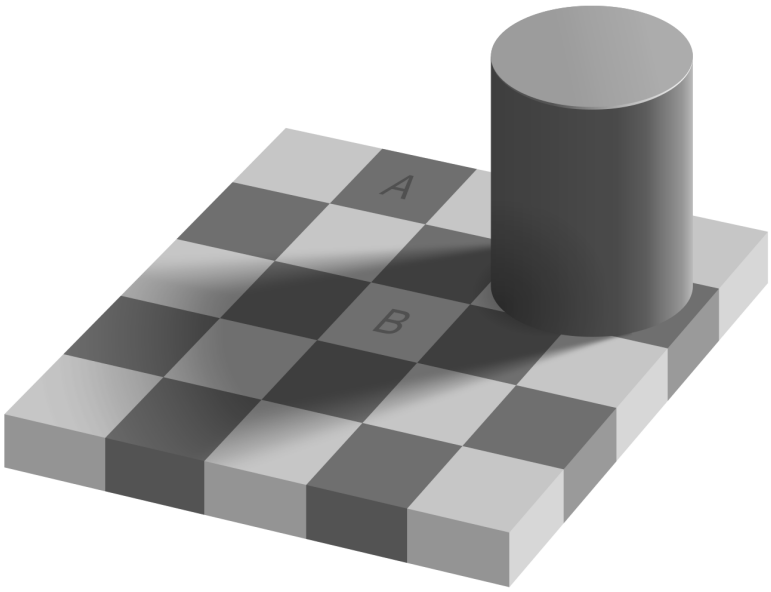
One possibility is that intuitions are not a distinct concept. It could be the case that intuitions are judgements or beliefs (Pust). Of course, the veracity of this proposal in turn depends upon the definitions of "judgements" and "beliefs". To avoid an unnecessary diversion into endless definitions, suffice it to say that common-sense conceptions will do. In this view, to have the intuition that something is true is exactly the same as believing it is true. So, if one believes that Citizen Kane is a great movie, that $2 + 2 = 4$, that capital punishment is wrong in all circumstances, and that the weather will be partly cloudy tomorrow then that is exactly the same as one having intuitions that those things are true. This means that intuitions have no unique properties

that beliefs lack and vice-versa. The two are totally indistinguishable in this view.

The idea that intuition is simply a synonym for belief or judgement is appealing because it makes the philosophy easier. All philosophical work on the nature of beliefs can be exported to intuitions. Besides its convenience, the view that intuitions are just beliefs passes a few sanity checks. If you have the intuition that capital punishment is wrong in all circumstances, then it makes sense to say that you believe it is wrong. If you later change your mind and find some circumstances where capital punishment is permissible, then it is plausible that you no longer hold the intuition that it is wrong in all circumstances.

While equating intuitions and beliefs would be convenient, this view does not stand up to scrutiny, as it fails to accommodate familiar mental phenomena. For one, it does not permit the notion of a counter-intuitive belief. It is quite possible to believe something after rational investigation, all the while holding the intuition that the proposition in question is false (Pust). Optical illusions are a classic example.

In the checker shadow illusion, two labeled squares on an image of a checkerboard appear to be two different shades of grey. In fact, the two shades are identical. Even after verifying the fact in a variety of ways, be it a photometer or literally cutting out the squares from the image and placing them right next to one another, one can't help but see two distinct shades in the original image. In this case, the belief that the two shades are not



Checker shadow illusion (Adelson). The squares labeled A and B are the same shade of gray.

the same may indeed be identical to the intuition upon first examining the image. But after investigation, the belief changes to the shades are the same color and the intuition remains that they are different. So, it is possible for your beliefs to change while your intuition remains the same. Beliefs can differ from intuitions in ethical cases, as well. It is possible, for example, for someone to commit to vegetarianism out of ethical concerns without intuitively finding eating animals problematic or off-putting. This hypothetical individual holds the belief that eating animals is ethically wrong without having the intuition that eating animals is wrong. They do not intuitively feel that eating animals is repulsive. Instead, they could believe it is wrong due to reasoning about their diet contributing to global warming and inhumane

factory farming.

Before moving on to the next view of intuition, it is important to understand the nature of the above argument and the ones to come. The above did not empirically prove that intuitions are not the same as judgements or beliefs. After all, these words refer to concepts that we associate with particular mental states. It is impossible to objectively measure or define the notion of a belief in the purely physical realm of atoms, such that a science experiment could sort it all out. It is true that a belief may be associated with the firing of specific neurons, but that is different from the phenomenon of experiencing the mental state of believing something. All this is to say that when trying to define intuition, the best strategy is to choose a definition that maps onto our phenomenological experiences. Equating belief and intuition fails that mapping.

2.2 Are intuitions dispositions to belief?

Another proposed definition of intuitions is that they are dispositions to believe in something. Under this view, you have the intuition that something is true if and only if you are disposed/inclined to believe it is true (Pust). Peter van Inwagen describes intuitions as "the tendencies that make certain beliefs attractive to us, that 'move' us in the direction of accepting certain propositions without taking us all the way to acceptance" (309). This view of intuition is an improvement over the previous one since it allows for belief and intuition to differ. One can be "moved" in the direction of accepting

something, yet still believe something else entirely. The definition is also successful in retaining some relationship between intuition and belief, which seems appropriate, as intuitions do seem to be the starting point from which people form their beliefs.

This definition is an improvement but suffers from being too broad in that it includes things which we would not generally consider to be intuitions as fulfilling the necessary criteria (Pust). Consider you stumble across a secret wish list that a young family member made for the holidays. At the top of the list, the child wrote what they most want as a gift. The wish list is not dated, so may be old, and the child in question is historically fickle. You do not go so far as to believe that the child still prefers the gift at the top of the list, but you are definitely disposed to believing so. According to the disposition to belief definition of intuition, that disposition qualifies as an intuition. Yet, this type of reasoning and mental state on your part is not what we mean when we refer to intuition. This definition of intuition is overly broad because any reason at all to have a belief would count as "moving" one to have that belief and consequently qualify as an intuition.

This definition is also too broad because it includes dispositions that have never, in fact, occurred (Pust). One may be disposed to believing something that they never considered. For instance, a friend of mine might be disposed to believe in a conspiracy theory, but unless someone tells them or they otherwise learn about it, then it does not make sense to say they have an intuition about it. Yet, this unconscious and never-accessed disposition

would qualify as an intuition under this definition. This does not make sense because intuitions are conscious mental states that are experienced. We may be unconscious or unaware of how an intuition came to be, but we are conscious of the intuition when we experience it. Counting hypothetical mental states that have never occurred as intuitions does not capture what is meant by the term.

2.3 Are intuitions emotions?

Some academics, particularly moral psychologists, view moral intuitions as a type of emotion (Clavien and FitzGerald 3). When I discuss the origins of intuitions in the next chapter, I point out a theory that focuses on emotions like compassion, anger, guilt, disgust, shame, fear, and pride (Haidt et al. 58). The authors of the theory call these "moral emotions" and point out that they occur during a moral evaluation of some action (Haidt et al. 13). The paper does not go into detail distinguishing moral intuitions, moral opinions, moral judgements, and moral emotions, all of which it mentions, but it does at least associate moral intuitions with moral emotions, if not equating them. One account is that moral intuitions arise from moral emotions. For example, a situation might trigger disgust which can be interpreted as a moral intuition itself or causing the moral intuition that the situation is ethically wrong. Thinking of moral intuitions as emotions does capture some aspects of the phenomenon of having an intuition. Intuitions, like emotions, are spontaneously generated in response to some stimulus. They come au-

tomatically and are hard to control, also like emotions. This makes sense considering the psychological underpinnings of intuitions, explained in the next section.

However, there are some differences between emotions and intuitions. First, it is possible to experience emotions without any ethical sentiment attached. Someone who is happy enjoying a sunny day is not necessarily making any ethical evaluations or polling their moral intuitions. Likewise, there are intuitions that have no associated moral emotion. Clavien and FitzGerald explain, "One can have a strong moral intuition about how to fairly distribute a cake among children at a party without feeling particularly emotional about the matter" (3-4). I think it's helpful to think of moral intuitions as sometimes associated with emotions and sharing similar characteristics, while recognizing they are distinct phenomena.

2.4 Intuitions are a unique mental state

Even after taking stock of the above definitions' defects, it is clear that an intuition must be connected to beliefs or dispositions to believe somehow. Intuitions make up a large part of our everyday judgements and moral beliefs. I believe the most promising view of intuitions is that they are a unique occurrent mental state, best thought of as "seemings" (Pust). In this sense, having the intuition that something is true is the same as having the conscious mental state that makes it "seem" true, but not on the basis of conscious reasoning, as in the wish list case. That last criterion preserves the

gut feeling aspects of intuitions that we experience. Henry Sidgwick used the term "spontaneous unreflected judgements" (qtd. in Appiah 75), which comes close to capturing this definition. My only qualm is the term "judgement", which is synonymous with "belief", and therefore problematic for the reasons explained in the earlier section. Admittedly, an intuition that makes a belief "seem" true is very similar to a disposition to that belief. The critical difference is not in the exact word "seeming", but in the conscious mental state and gut feeling it captures that mere disposition does not.

This view of intuition allows for counter-intuitive beliefs and beliefs that do not match intuitions, more broadly. It only includes conscious mental states but precludes conscious reasoning. Another boon to this definition is that it aligns with the psychological understanding of intuitions. Psychologists Keith Stanovich and Richard West coined the terms System 1 and System 2 to refer to two systems of the mind, the first being the system responsible for intuitions (Kahneman 20). Daniel Kahneman writes, "System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control" (20) and "System 1 continuously generates suggestions for System 2: impressions, intuitions, intentions, and feelings" (24). This description of System 1 aligns with Sidgwick's characterization of intuitions as "spontaneous" and "unreflected". It also explains why intuitions share some aforementioned characteristics with emotions. Here is a list of examples that are generated by System 1 thinking:

> Detect that one object is more distant than another.
> Orient to the source of a sudden sound.
> Complete the phrase "bread and . . . "
> Make a "disgust face" when shown a horrible picture.
> Detect hostility in a voice.
> Answer to 2 + 2 = ?
> Read words on large billboards.
> Drive a car on an empty road.
> Find a strong move in chess (if you are a chess master).
> Understand simple sentences.
> Recognize that "meek and tidy soul with a passion for detail" resembles an occupational stereotype. (Kahneman 21)

2.5 How intuitions fit into the bigger picture

As discussed previously, intuitions are distinct from beliefs. That being said, philosophers often use the word "intuitions" to refer to people's ethical beliefs, which can make talking about intuitions and beliefs confusing.

> Normative ethicists tend to think of moral intuitions as a kind of judgment, belief, conviction, or proposition (Cohen 2008; DePaul 2006; Kamm 2006; McMahan 2013; Rawls 1971). Most ethicists do not give a particularly detailed descriptive account of intuition because they are more interested in the practice of normative ethics than in moral psychology. The moral judgments that are thought deserving of the label "moral intuition" are characterized in different ways, e.g., as strong (Sinnott-Armstrong 2008), self-evident (Audi 2009; Shafer-Landau 2005), or as the result of long and considered reflection (Kamm 2006). On these views, having the moral intuition that abortion is wrong involves having a strong, self-evident, or considered judgment that abortion is wrong. (Clavien and FitzGerald 3)

Most of the time when philosophers do this, people's ethical intuitions

and beliefs are the same, so the distinction is not that important. But for our purposes, it is important to recognize that there is a distinction since we are going to talk about cases where ethical intuitions and beliefs differ. So, just be aware that when a philosopher says that someone has the ethical intuition that A is true, they often mean someone has the ethical intuition and the ethical belief that A is true. (By "ethical intuition", I do not mean that having the intuition is itself ethical, but rather that the subject matter of the intuition has to do with ethics.) Or, they really do not think there is a difference between intuitions and beliefs, but that is not the clearest understanding as previously discussed.

To ultimately understand cases where one's ethical intuitions and beliefs differ, think of intuitions as subservient to beliefs. Intuitions are a starting point, but beliefs are the final call, in other words what one actually thinks about something. So, when people disagree about something, they are disagreeing about their beliefs. They might very well also have differing intuitions, but that is another level removed from their disagreement. People's beliefs can be informed by many different sources and intuitions are just one of those sources, albeit an important one.

A helpful analogy to understanding how intuitions relate to beliefs is the five senses: sight, hearing, taste, smell, and touch. It is easy to think of what you sensorily experience as giving you a direct view to what actually exists, but a more accurate view of your senses is that they contribute to your beliefs about what exists. René Descartes had this line of thought in one of

his most famous thought experiments, where he wrote, "Thus, because our senses sometimes deceive us, I wanted to suppose that nothing was exactly as they led us to imagine" (18). Without thinking about it much, you might think that whatever you see exists. But after recalling optical illusions like the checkered shadow, you might realize, as Descartes did, that things are not always as they appear to be. Sometimes we swear that we hear a knock on the door or distant footsteps, but we end up mistaken. We heavily rely on our senses to gather data that we use to form beliefs about the physical world, but those beliefs can have other influences. If you measure the shades of grey in the checkered shadow illusion with a spectrometer, find that they are indeed identical, you can come to hold that belief despite your eyes telling you otherwise. Intuitions can be thought of similarly. We rely on them heavily to inform our beliefs, and most of the time they do equal our beliefs. Kahneman writes, "If endorsed by System 2, impressions and intuitions turn into beliefs. When all goes smoothly, which is most of the time, System 2 adopts the suggestions of System 1 with little or no modification. You generally believe your impressions and act on your desires and that is fine – usually" (24). But it is important to recognize intuition as a distinct phenomenon that informs belief.

2.6 Summary

Intuition is a unique mental phenomenon distinct from belief, disposition to belief, and emotion. Intuitions can be thought of as spontaneous, un-

reflected, and occurrent seemings generated by System 1 thinking. Beliefs are informed by intuitions, but also from other sources. While intuitions and beliefs are mostly in alignment, this need not be the case. Now having a thorough understanding of what I mean by the word "intuitions", it is tempting to jump right to discussing the heart of this paper, which is their reliability. We know what intuitions are, but are they useful at all? But before discussing reliability, it is necessary to understand where intuitions come from. The origins of this mental state could have implications on its reliability.

3 Where do moral intuitions come from?

Now that I have outlined a view of what intuitions are, it next makes sense to investigate where they come from. It is plausible that knowing the source of intuitions can gives clues as to their reliability, in the same way that the source of any data impacts its veracity. This inquiry is very much related to the philosophy of intuitions, but the question is purely an empirical one. Suggested answers come from anthropology, biology, psychology, and other empirical sciences, rather than philosophy.

3.1 Evolutionary explanations

It is widely accepted at this point that many of our moral intuitions are owed to evolution. This fact explains some of what might be called moral

quirks, like that we care so much more about those around us and whose names we know, than of strangers far away. Peter Singer writes,

> Much as we may regret it, most human beings lack a general feeling of benevolence for the strangers we pass in the street. In evolutionary terms, when we consider the species as a whole, the unit of selection is too large for natural selection to have much impact. Despite the picture books we had as children, early human life was not, by and large, a struggle for survival between humans and sabre-tooth tigers. It was much more often a struggle for survival between different human beings. There is no evolutionary advantage in concern for others simply because they are members of our species. In contrast to the selection of individual organisms within the species, which is going on all the time, selection between different species happens too slowly and too rarely to play much of a role in evolution. ("Ethics and Intuitions" 334-335).

It is not self-evident that we *should* care for all human beings equally, no matter how well we know them. The point is not whether or not this common moral intuition is correct. The point is that it exists and is attributable to evolution. Singer also describes how intuitions of reciprocity may have evolved.

> Many features of human morality could have grown out of simple reciprocal practices such as the mutual removal of parasites from awkward places. Suppose I want to have the lice in my hair picked out and I am willing in return to remove lice from someone else's hair. I must, however, choose my partner carefully. If I help everyone indiscriminately, I will find myself delousing others without getting my own lice removed. To avoid this, I must learn to distinguish between those who return favors and those who

> do not. In making this distinction, I am separating reciprocators and nonreciprocators and, in the process, developing crude notions of fairness and of cheating. I will strengthen my links with those who reciprocate, and bonds of friendship and loyalty, with a consequent sense of obligation to assist, will result. ("Ethics and Intuitions" 336)

Humans and their ancient ancestors relied on social collaboration and tribalism to survive so it makes sense that our moral intuitions tend to reflect that.

3.2 Cultural explanations

While evolution certainly plays a role, it is not the whole picture. After all, humans largely share the same evolved traits, yet moral intuitions differ widely. Culture also plays a role. A common view is that basic notions like reciprocity, which were universally necessary for survival, underpin all moral intuitions but some cultures have codified them differently.

> The basic rule of reciprocity, which includes the ability to detect cheats and the sense of indignation required to exclude them, is natural in the sense that it has evolved, is part of our biological nature, and is something we share with our closer nonhuman relatives. But the more detailed rules of justice typical of human, language-using societies are refinements on the instinctive sense of reciprocity, and so may be considered artificial.
>
> Our biology does not prescribe the specific forms our morality takes. There are cultural variations in human morality, as even Herodotus knew. Nevertheless, it seems likely that all these different forms are the outgrowth of behavior that exists in social animals, and is the result of the usual evolutionary processes of

> natural selection. Morality is a natural phenomenon. No myths are required to explain its existence. ("Ethics and Intuitions" 337)

So, both culture and evolution are responsible for the origins of our moral intuitions. I believe the best-supported theory that takes both of these roles into account is Moral Foundations Theory.

3.3 Moral Foundations Theory

Jonathan Haidt popularized Moral Foundations Theory based on research he conducted with other psychologists and anthropologists. Their goal was to identify the basic building blocks of morality across cultures (Haidt et al. 4). By building blocks, the researchers mean fundamental moral values. A classical utilitarian, for example, would identify pleasure/pain as the only building block since it is the only moral value the theory takes into account (Sinnott-Armstrong). It is important to note, however, that utilitarianism is a normative ethical theory. This means the theory purports to be true and gives prescriptive advice. In contrast, the researchers emphasize "that our project is descriptive, not normative. We are not trying to say who or what is morally right or good. We are simply trying to analyze an important aspect of human social life. . . . We think it would be helpful for social psychologists, policy makers, and citizens more generally to have a language in which they can describe and understand moralities that are not their own. We think a pluralistic approach is necessary for this descriptive project"

(Haidt et al. 5). This means that the researchers were engaged on a purely anthropological and psychological fact-finding mission, in an attempt to find the building blocks of moral intuitions. It does not mean that whatever building blocks they purported to find are good or bad building blocks to build an ethical system from. The fact that the researchers ended up with a pluralistic approach only indicates that cultures have a plurality of values. It does not indicate that a "correct" ethical theory, if such a thing exists, is necessarily pluralistic.

Moral Foundations Theory integrates elements of both evolutionary and cultural explanations of ethical development (Haidt et al. 7). These elements make up the four main claims of MFT:

1. "Nativism - There is a 'first draft' of the moral mind" that is innate (Haidt et al. 7)

2. "Cultural learning - The first draft gets edited during development within a particular culture." (Haidt et al. 9)

3. "Intuitionism - Intuitions come first, strategic reasoning second." (Haidt et al. 10)

4. "Pluralism - There were many recurrent social challenges, so there are many moral foundations." (Haidt et al. 12)

The researchers identify five moral foundations that are culturally widespread but vary in importance between cultures (Haidt et al. 16, 38): care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation (Haidt et al. 12-14). They do not claim that these five are the only possible

moral foundations, but that the empirical evidence most strongly supports these. Further research is needed to evaluate other possible moral foundations (Haidt et al. 12). According to the theory, our moral intuitions are caused by violations of these foundations (Haidt et al. 37). For examples of how violating a foundation can evoke moral intuitions, consider the sanctity/degradation foundation. The authors write,

> Hominid history includes several turns that exposed our ancestors to greater risks from pathogens and parasites, for example: leaving the trees behind and living on the ground; living in larger and denser groups; and shifting to a more omnivorous diet, including more meat, some of which was scavenged. (Haidt et al. 14)

This is the evolutionary account for how the foundation came to be. Recall that evolutionary accounts must specifically be rooted in an adaptive advantage that promotes survival. The evolutionary mechanism proposed that does so by addressing the disease problem is the emotion of disgust.

> The emotion of disgust is widely thought to be an adaptation to that powerful adaptive challenge (Oaten, Stevenson Case, 2009; Rozin, Haidt, McCauley, 2008). Individuals whose minds were structured in advance of experience to develop a more effective "behavioral immune system" (Schaller Park, 2011) likely had an advantage over individuals who had to make each decision based purely on the sensory properties of potential foods, friends, and mates. (Haidt et al. 14)

The theory is that humans evolved to have the emotion of disgust in

order to stay disease-free. Furthermore, cultures have developed a sanctity/degradation moral foundation based on this phenomenon.

> Disgust and the behavioral immune system have come to undergird a variety of moral reactions, e.g., to immigrants and sexual deviants (Faulkner, Schaller, Park, Duncan, 2004; Navarrete Fessler, 2006; Rozin, Haidt McCauley, 2008). People who treat their bodies as temples are praised in some cultures for the virtues of temperance and chastity. (Haidt et al. 14)

The idea is that situations seen to violate sanctity can elicit disgust by triggering the parts of our brain that were evolved for the purpose of staying clear of diseases. Since this is an evolutionary feature common to all humans, violations of the sanctity foundation are culturally widespread. That being said, violations of this foundation are viewed as worse in some cultures than in others. "Specifically, participants in Eastern cultures (South Asia, East Asia, and Southeast Asia) expressed slightly greater Loyalty-and Sanctity-related moral concerns than did participants in Western cultures (United States, Canada, United Kingdom, and other Western European countries), which are consistent with established cultural differences in collectivism (Triandis, 1995) and the role of purity concerns in daily life and religious practices (Shweder et al., 1997)." (Haidt et al. 26)

I have mostly pointed to evolutionary accounts of foundation development, but Moral Foundations Theory also has room for cultural explanations. It was found that cultures in areas of greater historical sociopolitical turmoil were more likely to adhere to the Loyalty, Sanctity, and Authority founda-

tions. The suggested reasoning is that such cultures view their surrounding social world as dangerous and threatening and so are drawn to moral foundations that reflect political conservatism (Haidt et al. 27). In fact, the difference of which moral foundations are most closely adhered to by liberals and conservatives is consistent within cultures globally. "Graham and colleagues (2011) found that across the 12 world regions for which data were available, liberals consistently valued Care and Fairness concerns more than conservatives, whereas conservatives consistently valued Loyalty, Authority, and Sanctity more than liberals." (Haidt et al. 27).

The main takeaway of Moral Foundations Theory is that our moral intuitions are responses to violations of moral foundations and they often manifest in emotional responses like disgust.

3.4 Backpropagation

If you recall the previous chapter, I wrote that intuitions informed beliefs. Our System 1 mode of thinking generates intuitions, which we can use to form beliefs. The rest of this chapter has explored why System 1 exists and the reasons it generates intuitions the way it does. There were evolutionary and cultural explanations. In addition to intuitions informing beliefs, however, it is also possible for beliefs to inform intuitions. I will refer to this as backpropagation since the direction of influence, i.e. from belief to intuition, is opposite the usual case. This is another account of how someone could come to hold an intuition.

Consider a non-ethical intuition first. "For instance, young children in the process of learning basic mathematically may not find it compelling that $2 + 2 = 4$. But after having become familiar with calculation and understood the underlying logic of it, usually by physically adding objects together until the pattern 'clicks,' this type of calculation is processed in a Type 1 manner; whenever one asks them the sum of 2 + 2, the answer comes quickly and fluently to their mind" (Clavien and FitzGerald 8). In this case, it was not initially true for the young child that the System 1 part of their brain generated the intuition that $2 + 2 = 4$, which they then used to form the belief that $2 + 2 = 4$. It happened the other way around. Only after getting a grasp on basic arithmetic, does the intuitionist System 1 thinking come first. Pop quiz! What is 12 squared? Your System 1 should have blurted out the intuition 144 into your head, which you then used to form the belief that that is the correct answer. You likely did not do the actual arithmetic problem, either in your head or with pen and paper. You could do that to check your intuition, but it is the case that the intuition comes first. But the intuition only comes first after first having gone through the work in childhood of explicitly learning and thinking about arithmetic.

The same idea applies to ethical intuitions and beliefs. Recall the example of the vegetarian who is not intuitively opposed to eating animals, but still believes that it is wrong. They did not come to their belief by relying on their intuition, but by deliberative reasoning. For a while, their intuitions and beliefs are out of alignment. Over time, however, I would suspect that

their belief that eating animals is wrong would eventually backpropagate to their intuitions. In a few years, the vegetarian may very well be intuitively opposed, perhaps associated with the emotion of disgust, to eating animals.

Understanding a possible reason why backpropagation occurs is rooted in System 1 and System 2. System 1 relies on heuristics and automatic responses. Your brain would like to offload as much work onto System 1 as possible in order to be most efficient, as System 1 is running all the time (Kahneman 24-25). But, if System 1 continually generates intuitions that are out of date, so to speak, then System 2 has to kick in. If the intuition were aligned with the belief, then System 2 would not be needed. Therefore, the brain is incentivized to get the intuitions and beliefs back in alignment and it could do this by backpropagating the belief to System 1. While we know that backpropagation does occur, there is scarce psychological research as to the exact mechanism by which it takes place. The System 1 and System 2 account is one plausible way.

The existence of backpropagation means that the influence between intuitions and belief is actually bi-directional. This means it is possible to have an intuition that A is true that informs a belief that A is true. However, you later come to believe through means other than intuition that A is actually false. For a while, your intuition and beliefs mismatch. Eventually though, your belief backpropagates such that your intuition is now also that A is false. The intuition and belief are once again in alignment. To add another layer of complexity, it could then be the case that a separate belief, that B is true,

actually depended upon the intuition that A is true. Stated in propositional logic, B is true if and only if A is true and some other propositions are also true. Your new intuition that A is false causes you to give up your belief in B, now believing that B is false. This series of examples is quite abstract, but the main point is that the bidirectional influence between intuitions and beliefs can have ripple effects to associated intuitions and beliefs.

3.5 Summary

The origins of intuitions are both evolutionary and cultural. Having established an overview of where intuitions come from, we can move on to reasons to be skeptical of their reliability. Keep the different accounts of intuitions' origins in mind this section since some of the arguments against reliability depend upon where intuitions come from.

4 Skepticism about the reliability of moral intuitions

The second chapter defined what I mean when I refer to moral intuitions. The third chapter explored origins of moral intuitions and found both evolutionary and cultural explanations. But the mere fact that intuitions are rooted in evolution or culture does not mean that they are reliable or unreliable. By reliable, I mean track the moral truth, assuming such a thing

exists. This chapter reviews arguments that are skeptical of the reliability of moral intuitions. They claim that moral intuitions have little to no bearing on what is actually moral. So, according to these arguments, if you have the intuition that it is wrong to do something, then that should actually have very little sway in forming the belief that it is right or wrong to do that thing. To concretize, just because one has the intuition that stealing is wrong has no bearing on whether stealing is actually wrong, under this view. You can substitute stealing for any action.

4.1 Burden of evidence is on reliability

The simplest, though admittedly not the strongest, skeptical argument about the reliability of intuitions is simply that the burden of evidence must be on demonstrating reliability, not unreliability. There is no reason to assume that moral intuitions are reliable by default. There does not exist some obvious causal mechanism that directly relates having a moral intuition with the truthfulness of the intuition. Hence, those who want to defend the reliability of intuitions must be the first ones to bring evidence. Lacking that, moral intuitions should be considered unreliable.

4.2 Evolutionary debunking

One avenue for casting doubt on the reliability of moral intuitions is a class of arguments known as evolutionary debunking arguments. The basic

premise of these arguments is that moral intuitions are based on evolutionary adaptations. While the previous section showed that some intuitions have more of a cultural explanation than an evolutionary one, it can be granted that a large part of moral intuitions do have some basis in evolution. The next premise is that there is no reason to expect evolution's goals to align with what is moral. Evolution is singularly focused on promoting survival through the mechanism of natural selection. Singer writes, "The direction of evolution neither follows, nor has any necessary connection with, the path of moral progress. 'More evolved' does not mean 'better'. . . . So, while I have claimed that evolutionary theory explains much of common morality, including the central role of duties to our kin, and of duties related to reciprocity, I do not claim that this justifies these elements of common morality" (343). The basic question is, if our moral intuitions are rooted in evolution, and evolution has nothing to do with morality, then why should we expect our moral intuitions to be reliable?

4.3 Social Intuitionist Model

The Social Intuitionist Model (SIM) does not directly attack the reliability of moral intuitions, but rather casts doubt on the legitimacy of moral reasoning more generally. Jonathan Haidt formulated SIM as a precursor to Moral Foundations Theory (Haidt et al. 11). "SIM proposed that moral evaluations generally occur rapidly and automatically, products of relatively effortless, associative, heuristic processing that psychologists now refer to as

System 1 thinking . . . Moral evaluation, on this view, is more a product of the gut than the head, bearing a closer resemblance to aesthetic judgement than principle-based reasoning" (Haidt et al. 11). Haidt et al. are making a descriptive claim about the nature of moral evaluation, but if they are correct, then the legitimacy of moral reasoning does not bode well. Even defenders of the reliability of moral intuitions would not go so far as to defend all moral intuitions in all cases. They would likely acknowledge that some cases require deliberative thinking to override some intuitions. Yet, Haidt et al. believe this rarely happens. Under SIM, the primary reason the deliberative thinking of System 2 is invoked is to fulfill "social requirements to explain, defend and justify our intuitive moral reactions to others. This notion that moral reasoning is done primarily for socially strategic purposes rather than to discover the honest truth about who did what to whom, and by what standard that action should be evaluated, is the crucial 'social' aspect of the Social Intuitionist Model. . . . We reason mostly so that we can support our judgements if called upon by others to do so" (Haidt et al. 11). This explanation of moral reasoning, if correct, makes a mockery of the notion of truth-seeking moral deliberation and reliable intuitions.

Worse still for the reliability of moral intuitions, is how easily they are manipulated by capitalizing on moral emotions like disgust.

> Building on the finding that conservatives tend to moralize Sanctity concerns more than liberals (Graham, et al., 2009), Helzer and Pizarro (2011) reported two experiments in which subtle reminders of physical purity – standing by a hand sanitizer and us-

> ing hand wipes – led participants to report being more politically conservative and more disapproving of sexual purity violations, like incest or masturbation. Similarly, Inbar, Pizarro, and Bloom (2011) found that experimental inductions of disgust [such as bad smells] led participants to report more negative attitudes towards gay men but not towards lesbians or other outgroups. (Haidt et al. 24)

If simply being exposed to bad smells or a dirty environment is enough to influence a moral intuition, then how could we sincerely rely upon such intuition?

4.4 Cognitive diversity

Stephen Stitch raises the "cognitive diversity" objection. Simply put, the cognitive diversity objection is "about different ways of thinking . . . and the problem of choosing among them" (Stitch). The fact of the matter is that almost all intuition, be it moral, epistemic, or otherwise, is far from uniform. Consider moral intuitions held throughout history. The prevailing moral attitudes today differ from those held only a few decades ago, let alone centuries or millennia. It is not uncommon for people to look back aghast at what their ancestors considered moral or immoral, only for the next generation to do the same to them. There is no reason to believe that our current time period is any different; our descendants may look back horrified at some of our intuitions. In fact, an individual's moral intuitions may even evolve over the course of their lifetime. Diversity in intuition also exists across axes other than time, such as geographically and culturally.

There exists a wide range of intuitions among different cultures, all held in 2016.

The problem with cognitive diversity is that intuitions do not command authority over one another. Your say is as good as mine, as the saying goes. As Mitchell writes, "The problem is not merely that people disagree, but that their differing intuitions have the same authority. The most our intuitions can do, it seems, is tell us about ourselves and our own ways of thinking, not about the facts they're supposedly 'about'" (Mitchell). Consider the (in)famous trolley problem thought experiment. "Beginning with Philippa Foot and Judith Thomson, philosophers have invited readers to imagine that a trolley is speeding down a track. Unimpeded, the trolley will hit five people ahead of it, killing them, but an innocent person nearby could stop it. In one version, she could stop the trolley and save the five people by pulling a lever to divert it to another track, but this would kill one person who happens to be on that track. In another, she can only stop the trolley from killing the five by pushing someone off a bridge into the trolley's path" (Mitchell). What is to be done if, when faced with the trolley problem, your intuition tells you that pulling the lever is morally necessary while at the same time my intuition tells me that pulling the lever is morally inexcusable? The cognitive diversity argument is that since there is no way to arbitrate whose intuitions carry authority, we cannot rely on either of these intuitions to tell us anything about ethics.

There are two groups of researches known as radical experimentalists,

who argue against using intuition in philosophy (Liao 3). The first group, made of Jonathan Weinberg, Shaun Nichols, and Stephen Stich (abbreviated as WNS), tries to bolster the cognitive diversity objection by "show[ing] that intuitions about certain cases, which philosophers have taken for granted as being uniform, in fact vary according to factors such as cultural and educational background" (Liao 3). WNS hope to prove intuitions unreliable by demonstrating cognitive diversity to be a problem in many philosophical issues, even those considered settled or "taken for granted", rather than a rare problem. Before delving into the details of the radical experimentalists' arguments, I want to note that the examples they use are from philosophical fields other than ethics. This does not necessarily undermine their claims since intuitions are used largely the same way in ethics as in other fields of philosophy. An important question in epistemology, to give an example, is defining what knowledge is. For a long time, the answer was thought to be justified true belief (Ichikawa and Steup). In 1963, Edmund Gettier proposed a few counterexamples to this theory. In one example, Smith justifiably believes that Jones owns a Ford. Through disjunction, Smith justifiably believes that Jones owns a Ford or Brown is in Barcelona, without actually knowing where Brown is at all. As it turns out, Jones does not actually own a Ford and Brown, by pure coincidence, is in Barcelona. According to the justified true belief definition of knowledge, Smith's belief concerning Jones and Brown is knowledge. Our intuition, Gettier said, tells us that this is actually not knowledge – Smith just got lucky – and so the justified true

belief theory of knowledge is lacking (Ichikawa and Steup). The question WNS address is if using intuition to support or, as Gettier did, to oppose a theory is justified.

WNS pose many hypothetical examples to survey participants and noted the participants' differing intuitive responses to the same examples in order to demonstrate cognitive diversity. Their main point was that differences in social, economic, cultural, and educational backgrounds create differing intuitive responses to the same cases. One example was the Charles case: "One day Charles is suddenly knocked out by a falling rock, and his brain becomes re–wired so that he is always absolutely right whenever he estimates the temperature where he is. Charles is completely unaware that his brain has been altered in this way. A few weeks later, this brain re–wiring leads him to believe that it is 71 degrees in his room. Apart from his estimation, he has no other reasons to think that it is 71 degrees. In fact, it is at that time 71 degrees in his room. Does Charles really know that it was 71 degrees in the room, or does he only believe it?" (Liao 6-7). It would be typical of an epistemic philosopher to use intuitions about whether or not Charles possessed knowledge as evidence either in support of or in opposition to a particular theory of knowledge. Even though this particular example is not a moral case, the general argument still applies. The issue with this reasoning, as WNS point out, is that "East Asian subjects are much more likely to deny knowledge than their Western classmates" (Liao 7). Neither East Asian nor Western intuition can make any legitimate claim to authority over the

other. So, the objection goes, a philosopher might think that the clear intuitive response might be "Charles has knowledge", while many people might intuitively believe otherwise (or vice-versa). This is especially problematic when it comes to cases that philosophers considered settled, or at least more clear-cut than cases like Charles'.

Recall Gettier cases, where philosophers mostly agreed that there was no knowledge and so the justified true belief theory had to be modified. WNS ask, "Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen, and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car. Does Bob really know that Jill drives an American car, or does he only believe it?" (Liao 7). This is a Gettier case, yet "WNS find that a large majority of W[esterner]s give the standard answer in the philosophical literature, namely, 'Only Believes,' but a majority of E[ast] A[sian]s say that Bob really knows" (Liao 8). This finding demonstrates the problems of using intuition as evidence for or against a theory. Since the majority of East Asians believe Bob does have knowledge, perhaps the Gettier objections do not contest the justified true belief theory of knowledge. To be sure, the case is not as clear cut as philosophers may have once thought.

4.5 Recently considered cases objection

The second group of radical experimentalists, Stacey Swain, Joshua Alexander and Jonathan Weinberg (abbreviated as SAW), offer a new argument altogether, "reveal[ing] that intuitions about cases that have been regarded as uniform can in fact vary according to what other cases have recently been considered" (Liao 3). SAW's experiments also aim to demonstrate intuitions as unreliable, but not by revealing cognitive diversity. Instead, they ask different groups of participants the same cases, but in different orders. The idea is that, if intuition were reliable and truth-tracking, then the order we consider cases should not affect our intuition about those cases. The order should be an irrelevant factor. Yet, SAW show that the order does have an influence on intuition. Recall the case of Charles, who believed it was 71 degrees in his room. Now consider the case of Dave: "Dave likes to play a game with flipping a coin. He sometimes gets a 'special feeling' that the next flip will come out heads. When he gets this 'special feeling', he is right about half the time, and wrong about half the time. Just before the next flip, Dave gets that 'special feeling', and the feeling leads him to believe that the coin will land heads. He flips the coin, and it does land heads" (Liao 9). Dave clearly does not know the coin will land heads, he just got lucky. SAW's research reflects this belief across socioeconomic and cultural backgrounds. However, "SAW find that the subjects' willingness to attribute knowledge to Charles in the [temperature] case vary depending on whether, and what, other cases are presented before it. In particular, SAW find that the subjects' willing-

ness to attribute knowledge to Charles increase after being presented with a case of non-knowledge (the Coinflip Case)" (Liao 10). SAW then go on to present a case of clear-knowledge. After listening to the clear-knowledge case, participants were less likely to consider Charles as having knowledge. SAW assert that since order, an irrelevant factor, does influence intuitions then they are unreliable to use as evidence.

4.6 Summary

Hopefully I outlined the arguments skeptical of the reliability of moral intuitions sufficiently well so that you do have some serious doubts that they could be reliable. The above arguments raise serious objections and from multiple points of view. In the next chapter, I will rebut these arguments and defend the reliability of intuitions.

5 Defenses of the reliability of moral intuitions

Despite all the arguments in the previous chapter that gave very good reasons to be skeptical of the reliability of moral intuitions, I believe that we can rely on moral intuitions as starting points for moral beliefs. Before giving an overview of what I think the landscape of reliability looks like, I want to address each of the arguments from the previous chapter individually.

5.1 Response: Burden of evidence is on reliability

I agree with the spirit of this skeptical argument, but there are indeed some reasons to believe that moral intuitions are reliable. One can empirically observe that, for the most part, our moral intuitions help us successfully navigate through life. We can typically collaborate with others, try to help people out when we can, and strive to live a good life. Much of this is based on our moral intuitions. If our moral intuitions did not track the moral truth at all, then it is surprising how good people's lives generally are. Imagine that our moral intuitions had no connection to the moral truth whatsoever and so are essentially random. I posit that the world would be in absolute chaos. I am not making the argument that our moral intuitions are perfect all the time. Far from it. My main point is that a glass half full view of reliability is more accurate than a glass half empty one. By this I mean that I think everyone could agree that some moral intuitions are not correct. At the same time, I think you could persuade people that our moral intuitions are at least better than random, given the state of the world.

5.2 Response: Evolutionary debunking

I take a similar approach responding to evolutionary debunking arguments. It is true that moral intuitions are heavily informed by adaptations made. It is also true that there is no reason to believe that the outcomes of natural selection would perfectly align with what is ethical. However, I

disagree with the stronger claim made by evolutionary debunkers, which is that natural selection does not align with ethics at all. After all, a large part of survival for humans is collaboration with others. "David Copp . . . has argued that . . . it is plausible to suppose that a moral psychology shaped by natural selection in social contexts would yield moral beliefs that track moral truths at least reasonably well as a first approximation. On such a view, moral truths are grounded in moral standards having to do with codes that would enable societies to meet their basic needs (continued existence, stable cooperation among members, internal harmony and peaceful relations with other societies)" (FitzPatrick). I share this view. As I said earlier, I do not expect natural selection to produce moral intuitions that are always reliable all of the time. But I find it implausible to think that natural selection's goals are totally antithetical to morality. After all, I think most would agree that even the furthering of one's own survival, which is at the heart of natural selection, is ethical in some sense. Just as before, I take a glass half full perspective. I do not think that moral intuitions are perfectly reliable in that they always correspond to the moral truth. However, I think there are reasons, evolutionary accounts among them, to believe that moral intuitions are reliable enough as a starting point to form accurate moral beliefs.

5.3 Response: Social Intuitionist Model

I do not dispute the empirical claims that Haidt et al. make about how most people engage in moral reasoning most of the time. I do unfortunately

find it plausible that people tend to use deliberative reasoning only to find justifications for their preexisting intuitions. It is also unfortunate that morally irrelevant factors like bad smells can influence our moral intuitions on totally unrelated cases. That being said, it does not mean that this is the only type of moral reasoning that is possible. Serious philosophical inquiry can still occur and genuinely use deliberative thinking to question intuitions, rather than simply rationalize them. While this may not be the predominant form of moral discourse, that is okay. We can be deliberative and questioning if we so choose.

5.4 Response: Cognitive diversity

I have not encountered any argument that explains how conflicting intuitions can be used to assess a philosophical theory. However, the cognitive diversity objection does not do as much work as one might think it does at first sight. While there indeed appears to be no basis to assess theories using conflicting intuitions, it does not follow that all intuitive assessments are unfounded. The objection relies on an argument by question: "Since disagreement is a possibility, why should we think intuitions track the truth?" (Mitchell). Simply because some people may be mistaken does not imply that everyone is. I admit the cognitive diversity objection raises the more difficult question of determining which intuition is correct, but it does nothing itself to show that neither could possibly be correct. This argument also ignores the additional possibility for there to be some set of intuitions that

are truth-tracking and universal among those that are not. It does not have to be the case that all intuitions are entirely correct, or all intuitions are entirely incorrect. There are some questions, like is it permissible to commit patricide for no reason, that have little to no cognitive diversity.

5.5 Response: Recently considered cases objection

The same response can be made to SAW, whose research shows order influences intuition. While intuitions in response to some cases, like Charles', changed depending on order, SAW themselves admitted that others did not, namely the clear non-knowledge coinflip case and the clear-knowledge case. S. Matthew Liao says, in response to the radical experimentalists, "But from the experimentalists' own, empirical methodology, the fact that these [unambiguous] cases exist do suggest that we can sometimes rely on our intuitions; or at the minimum, they suggest there is no empirical evidence to think that we could not rely on these intuitions" (Liao 12). Liao asserts that while cognitive diversity or change due to recent consideration of other cases may exist sometimes, there is a core set of intuitions that are truth-tracking and reliable. In response to the findings that all groups, irrespective of background or recent cases considered, did not find the coinflip case to be knowledge, Laio says, "[The radical experimentalists] fail to recognize that this is a large concession, in fact, to the extent that it undermines their claim that there is not a fixed set of intuitions about a particular thought-experiment to which we can appeal. Moreover, once it is granted that some intuitions can serve

as evidence, this opens up the possibility that other intuitions might also be part of this common core and might also be able to serve as evidence" (Liao 12).

Another response to the recently considered cases objection is conceding that this psychological fact makes our intuitions less reliable than they would be otherwise, but still maintaining that it is possible to form accurate beliefs. If someone is going through the deliberative process of forming a moral belief, they can take care to poll their intuition multiple times and over multiple days to minimize the chance of being affected by a recently considered case or some other priming mechanism.

5.6 Summary

None of the skeptical arguments against the reliability of intuitions is fatally flawed. They raise serious objections and some empirically show that there are unavoidable reasons that make intuitions less reliable than they would be otherwise. However, none of the arguments is successful in demonstrating that moral intuitions are so consistently unreliable that they cannot be used as a basis to form moral beliefs. Having finally established a clear picture of intuitions, their origins, and their reliability, the rest of the paper is dedicated to answering fundamental questions about how to use intuitions to form accurate beliefs, contest other intuitions, and craft theories.

6 Moral disagreement

The cognitive diversity objection pointed out that people often have differing moral intuitions about the same case. How should people settle moral disagreement considering there is no obvious way for one individual's intuition to have authority over another's? Remember that moral disagreement is firstly about differing beliefs. The underlying moral intuitions that informed the beliefs may also be in disagreement, but they need not be.

Disagreeing with someone about a moral issue is very different from disagreeing about an arithmetic problem. There is no calculator readily available to settle the dispute. And, also unlike arithmetic, there is no guarantee to a right answer at all. Yet, despite these hurdles, people do change their mind on moral matters. Consensus can sometimes be reached. How does this happen?

6.1 Empirical disagreement

Most moral disagreement can be separated into two parts: empirical and normative. Empirical has to do with the facts of the matter, i.e. what is objectively and observably true that is relevant to the disagreement. Normative is deciding what is to be done after taking stock of the empirical aspects of the situation. Say you witness a bank robber steal a bunch of money and escape. Later that day, you notice the criminal handing over wads of cash, but for a donation to a good charity. Do you call the police or stay silent?

The empirical aspects of the situation are the bank robbery and the donation to the charity. The normative aspect is what you choose to do given that information.

It turns out a lot of moral disagreement is really rooted in empirical disagreement. Sometimes those in disagreement are not even working from the same set of shared facts, so it is a small wonder they do not come to the same moral conclusion. You might think it is okay to let the bank robber go since the money is going to a good cause, while your friend might disagree and want to call the police. It turns out, however, that your friend has access to empirical information you did not. The charity is really not a good cause after all, but a front for a money laundering operation. Upon learning this from your friend, you agree to call the authorities. It turned out that there was no normative disagreement, simply an empirical one. It is important that all parties agree on the relevant facts before debating what to do. In this case, you and your friend had differing moral beliefs without having differing moral intuitions. It was the differing empirical information that accompanied the intuition in forming the belief that was the root of the disagreement.

6.2 Normative disagreement

But sometimes there is genuine moral disagreement when all facts are agreed upon. There are no empirical disputes, yet two people make different normative evaluations. This is a case where the moral intuitions are in conflict. What is to be done?

One persuasion technique is to offer a counter-example. Say you have an ethical disagreement with someone about what to do in a certain scenario. You can come up with a different scenario, even an entirely hypothetical one, that is identical to the original scenario in every relevant ethical aspect. To conceptualize this, imagine each possible ethical aspect of the scenario is represented on its own axis. Another name for ethical aspect is ethical value. There might be a pleasure axis, a dignity axis, an innocence axis, and so on. The point on each axis captures the information of how its ethical value pertains to the scenario in question. Some aspects are not so simple as a quantity along a number line, but take the simplified idea of an axis to encapsulate all relevant quantitative and qualitative information for a single aspect. These axes represent the empirical information about the scenario. The basis of your intuition should rest on these axes. A classic utilitarian would only care about the pleasure axis, for example. But whatever your personal ethics, if you consider two scenarios that share the same points along their ethical axes, you are compelled to believe the same thing in each case to be consistent. If you grant the two cases are identical in all possibly ethically relevant ways, there is no logical basis to discriminate between the two.

This axes idea is similar to Jonathan Dancy's concept of shape (Dancy 111). Instead of axes, it might be easier to view each situation as having its own unique shape that represents it. Two entirely different situations will have two entirely different shapes. In this case, the shape represents all the

empirical information about the situation.

So, if you find yourself in a normative disagreement, a prudent strategy is to offer a counter-example that is ethically identical to the one disagreed upon, but also where you suspect your debater's intuitions align with your own. If you can convince them that the two situations are ethically identical, then they will be logically compelled to change their mind about one of them. You hope that they change their mind about the original case, rather than the counter-example, in which case you would be in disagreement about both cases. It may be difficult or outright impossible to construct an ethically identical case, in which case your goal is to create one as close as possible to the original.

Consider this concrete example of this practice from Peter Singer's drowning child thought experiment. Singer believes that people who live well-enough in developed countries are morally obligated to direct some of their earnings to overseas humanitarian charities. Let us assume that you do not believe that you are morally obligated to do so. Perhaps you think it would be a good thing for someone to do, but nowhere near obligatory. Singer would respond by asking you to imagine that your route to work passes a shallow pond.

> One morning, I say to [you], you notice a child has fallen in and appears to be drowning. To wade in and pull the child out would be easy but it will mean that you get your clothes wet and muddy, and by the time you go home and change you will [be late to work] . . . do you have any obligation to rescue the child? ("Drowning

Child").

I hope that your moral intuition and belief is yes. In any case, it must be yes for this argument to succeed. This is the counter-example that polls your intuition so that you are set up to agree with Singer in the next part. Singer continues constructing the counter-example so that the shape or ethical axes more closely resemble the overseas charity case.

> Does it make a difference, I ask, that there are other people walking past the pond who would equally be able to rescue the child but are not doing so? No, [you] reply, the fact that others are not doing what they ought to do is no reason why I should not do what I ought to do. Once we are all clear about our obligations to rescue the drowning child in front of us, I ask: would it make any difference if the child were far away, in another country perhaps, but similarly in danger of death, and equally within your means to save, at no great cost – and absolutely no danger – to yourself? Virtually all agree that distance and nationality make no moral difference to the situation"("Drowning Child").

Singer then points out that you are essentially in this situation right now. There are children in danger of death well within your means of saving through a donation to an overseas charity. If your intuition throughout the drowning child case was that you are obligated to help, but your intuition for the overseas charity case is that you are not obligated to help, you are now forced to reconcile these views in order to have a consistent set of beliefs.

One option is to accept Singer's argument, change your belief that you are indeed obligated to donate to overseas charities, overriding your intuition that

that is not the case. If you do so, your newfound belief may backpropagate so that later you do intuitively find that you are obligated to donate.

If you do not accept Singer's argument, there are only two fundamental ways to maintain consistency. These ways are not tied to this particular moral situation, but are inherent in the counter-example intuition argument structure. The first is to say that you do not share the intuition of the counter-example, in this case Singer's intuitions that you are obligated to save the drowning child. The second option is to insist that the ethical axes' values or shape of the counter-example and the original example are not the same and that this difference is morally relevant. A specific example of this might be to cast doubt on the efficacy of the donation. You know that you will be successful in saving the drowning child, but you are not sure that the charity will get the money to those who need it. Whether this is a good response in this specific case is not the point; I just mean to illustrate the general technique one would have to employ in order to defeat the counter-example argument. If it were true that the charity was not effective in helping save children's lives, then it is permissible to not believe you are obligated to donate, even if you have the intuition that you are obligated to intervene in the drowning child case.

6.3 Summary

Moral disagreement is disagreement between moral beliefs. Moral beliefs do involve moral intuitions, but include other data like empirical information.

Until two people agree upon the facts of the matter, or shape, no progress can be made. If people agree about the shape but still have differing moral beliefs, then they have differing moral intuitions. One person can offer a counter-example substantially similar to the original case, aiming to elicit a moral intuition from the other person that would put the two in agreement. The person who is asked to consider the counter-example might find an empirical difference between the two cases that make them morally distinguishable. If there is no ethically relevant empirical difference and they admit they have the moral intuition that is in contradiction to their intuition of the original case, then they must change one of their beliefs in order to be consistent. I want to emphasize that they change their belief, not their intuition. One cannot voluntary change their intuition since it is a product of the automatic System 1. It may change via backpropagation over time, but this is not a voluntary process.

7 Intuitions and theory

I can now address one of the fundamental questions of the paper, which is how should intuitions and theory interact? As a reminder, it is commonplace in philosophy to judge ethical theories or frameworks based on how they align with moral intuitions. If a theory recommends a counter-intuitive action, then it is judged negatively. At the same time, however, there must be same cases where the theory recommends a more ethical action than intuition does,

for if not, then ethical theories have no purpose at all. Before addressing this contradiction directly, I want to draw a distinction between what I will call "idealized" and "practical" ethical frameworks.

7.1 Idealized and practical ethical frameworks

Ethical frameworks prescribe specific actions in specific circumstances. The frameworks claim that their suggested actions are ethically superior to alternative actions. One can divide frameworks that fit this description into two sets. The first are frameworks that purport to represent the moral truth and address all circumstances. I refer to these as idealized frameworks. Idealized frameworks not only prescribe advice, but insist that such advice is defined as ethically correct. Idealized frameworks, then, assume moral realism, which is the position that there are such things as objectively true moral facts (Sayre-McCord). To take an action not advised by the idealized framework is to do something ethically wrong, according to the framework.

The second set of frameworks are what I will call practical frameworks. Practical frameworks also prescribe advice, but do not claim that this advice is identical to the definition of what is ethical. Practical frameworks are more modest in their ambitions. They do not aim to define what is moral, but only give practical advice to achieve ethical outcomes. This does not mean practical frameworks necessarily deny moral realism. A practical framework is simply agnostic to what is actually ethically true, and gives advice that could adhere to a plurality of possibilities of what is ethically true.

To clarify this distinction in ethical frameworks, I will give an example of each. Classic utilitarianism is an idealized ethical framework that claims an act is morally right if and only if that action results in the greatest net pleasure for all (Sinnott-Armstrong). This framework is prescriptive since it advises you to act according to what will have the maximally beneficial consequences. At the same time, the framework purports to be true, meaning it offers a definition of what constitutes moral action. Maximizing Wealth Plus, on the other hand, is an example of a practical ethical framework. It advises actions that maximize "[t]he total amount of value produced over a certain time period. This includes the traditional measures of economic value in GDP statistics, but also includes measures of leisure time, household production, and environmental amenities, as summed up in a relevant measure of wealth" (Cowen 30). Cowen argues this is best accomplished by policies that maximize sustainable economic growth and individual action that accords with common-sense morality, i.e. hard work, looking after your family, and so on (23). The merit of Cowen's Maximizing Wealth Plus framework is irrelevant to this discussion; it is simply an example of a practical ethical theory. It is practical, rather than idealized, because it does not define Wealth Plus as identical to what is morally valuable. In contrast, classic utilitarianism does define the presence of pleasure and absence of pain as the only moral values. Cowen's framework gives practical advice to bring about a more ethical world, but is silent on what exactly is, at bottom, of moral value.

It is important to distinguish between idealized and practical ethical theories because they each have unique properties, which open them up to unique forms of critique. A poor critique of classic utilitarianism, and of any idealized framework, is that it is impractical. One might argue that it is too difficult to predict the consequences of actions, so basing an ethical framework on consequences is a bad idea. This criticism is poor because it mistakes classic utilitarianism for a practical theory, rather than an idealized one. Idealized frameworks exist to define what is, in fact, ethically true. There is no requirement for the ethical truth to be practical, simple, or undemanding. The difficulty of predicting consequences does not impact the veracity of consequences having moral relevance. It would be nice if, assuming moral realism is true, morality were practical, simple, and undemanding, but it need not be. Yet, these types of critiques are perfectly valid when levied against practical frameworks, such as Maximizing Wealth Plus. Cowen only gives practical advice for real-world action, so it makes sense to critique the framework if the advice is impractical.

Critiques that appeal to ethical intuitions pose more of a problem for idealized frameworks than for practical frameworks. If the advice of a practical framework conflicts with an intuition, one can defend the framework in two ways. The first is to argue that the intuition should be discarded in favor of the framework's suggested action. Idealized frameworks also have access to this defense against a conflicting appeal to intuition. But practical frameworks have a second defense, which is to simply defer to the intuition.

Since practical frameworks do not purport to define the moral truth, there is no problem with admitting intuitions may be better in some cases. The practical framework may be handy to follow in most cases, but not all. Idealized frameworks do not have this luxury of deference. Any intuition that conflicts with the advice of an idealized framework must necessarily be incorrect if the framework is to be true. The framework must give the correct advice in all cases, or else be flawed. When faced with a conflicting intuition, the framework's only chance of survival is to convince the individual to reject their intuition. If the intuition cannot be rejected, then the only option to maintain consistency is to modify the framework to accommodate the intuition or discard the framework entirely.

7.2 Reflective equilibrium

The question I started with at the beginning of this chapter has narrowed to how do you decide between rejecting an intuition or modifying/rejecting an idealized theory when the two conflict? The longstanding answer in modern philosophy is a process known as reflective equilibrium, formulated by John Rawls. Kwame Appiah captures the sentiment that reflective equilibrium is not such a satisfying answer. "What has made [reflective equilibrium] so durable is not that it has solved the difficulties it means to address but that the difficulties themselves have proved so durable . . . Indeed, one could be forgiven for thinking that reflective equilibrium is really another name for the problem, rather than a solution to it" (78). Reflective equilibrium is a fancy

phrase that simply means adjusting particular beliefs and principles until you have reached an "equilibrium", meaning a set of positions that are coherent with one another. The lack of specificity as to how one is supposed to choose what to adjust in order to achieve coherence makes the methodology quite useless.

Dietmar Hübner writes that Rawl's original conception of reflective equilibrium made sense only due to the specific context of the argument he was making in A Theory of Justice (1). Philosophers since Rawls have used reflective equilibrium outside of that context, such as something "conceived of as the proper adjustment of theoretical accounts (based on scholarly argument, systematic elaboration, etc.) and intuitive convictions (rooted in a more direct access to moral features, non-inferential evidence concerning ethical value, etc.)" (Hübner 8). This modern tradition is actually an instance of "historical dissonance" and a misapplication of Rawl's methodology (Hübner 8). Hübner points out "the initial assumption—that RE should hold between theory and intuition—is already flawed: this idea is neither supported by Rawls' original account of the concept in 'A Theory of Justice' nor apt to ensure its fruitful application" (8).

Using reflective equilibrium to adjudicate disagreements between moral intuitions and moral theories is a misapplication of Rawl's methodology. Even if it were not, the methodology is not specific enough to give satisfying answers to how one is supposed to achieve a coherent equilibrium.

7.3 Alternatives to reflective equilibrium

Unfortunately, I have not devised a foolproof methodology of my own to put millennia of ethical debate to rest. However, I do believe I can at least improve upon reflective equilibrium. When a moral intuition and idealized theory conflict, it is actually the case that two moral beliefs are in conflict. One might have the moral intuition that A is true, but an idealized theory says that A is false. This is actually the same scenario of moral disagreement that was discussed in the previous chapter. The same techniques can be applied here.

One can come up with a substantially similar situation in order to test their intuition. The situation should elicit the intuition that A is false, like the idealized theory. If one does have the intuition that A is false in the similar situation and does not find any morally relevant differences between the cases then one should adopt the moral belief advocated for by the theory, rejecting the initial intuition.

To test the theory, rather than your intuition, think of a substantially similar situation that elicits the intuition that A is true. Since the idealized theory is not a person to argue with, it will never change its mind. The best you can do is come up with a counter-example that elicits a moral intuition that you so strongly believe in and the theory does not, or vice-versa, that you conclude the theory must be modified. This is not a wholly satisfying answer, but it is the best one can do. I would suggest trying to find a strong intuition that the theory disagrees with, not one that simply seems incorrect

at first glance. Another strategy is to enlist a friend to play devil's advocate on behalf of the theory. They can suggest counter-examples to try and move your beliefs in favor of the theory.

There is a final tactic to determine the reliability of the intuition in question and it takes advantage of the empirical characteristics of System 1 thinking. System 1 tends to be reliable when it is has been disciplined by immediate and definitive feedback (Chabris and Simons 145). This is why chess players can develop better intuitions about good moves than doctors can develop about medical diagnoses. When a chess player plays a bad move, they are instantly punished for it. Medical treatment, on the other hand, takes place over a long period of time and outcomes are not systematically collected and analyzed. Applying this to moral intuitions, we should be most skeptical of intuitions that come from unusual scenarios, ones we do not experience on a day to day basis. The moral intuitions that allow us to get through our daily lives without spending energy in deep deliberative moral judgement are fine-tuned everyday by definitive and immediate feedback. We should expect moral intuitions about unfamiliar cases to be more unreliable.

8 Conclusion

Intuition is a distinct mental phenomenon that is a result of System 1 thinking. Intuitions are occurrent, spontaneous, automatic, involuntary, and

unreflected. Moral intuitions, in combination with empirical information, inform our moral beliefs. When there is moral disagreement, it is always the case that moral beliefs are in disagreement. Moral intuitions need not be in disagreement in such cases, though they can be.

Intuitions are products of evolution and culture. We know that not all intuitions will be reliable all of the time, but there is reason to believe some moral intuitions are reliable. If moral intuitions were entirely wrong, it is unlikely the state of the world would be this good. Evolution, while not selecting for moral truth, does select for self-preservation and social collaboration, which are important elements of ethics. Moral intuitions are reliable enough to use as a starting point from which one can use to form moral beliefs.

The best way to settle moral disagreement is to first agree upon the facts of the matter. After that, generate similar counter-examples with the aim of eliciting your interlocutor's intuition to match your own. If the counter-example is not substantially similar to the original case, then it doesn't matter what the intuition is.

Ethical theories can be divided into the idealized and the practical. Both types are prescriptive but only idealized theories purport to define the moral truth. As a result, differing intuitions pose a greater problem for idealized theories. Reflective equilibrium is not a robust methodology to resolve intuition in conflict with an idealized theory. There is no foolproof method of arbitration, but there are psychological heuristics that have to do with

reliability. The techniques to settle moral disagreement can also be applied to test the intuition and theory.

Works Cited

Adelson, Edward H. "Checkershadow Illusion." *Perceptual Science Group @ MIT*, persci.mit.edu/gallery/checkershadow.

Appiah, Anthony. *Experiments in Ethics*. Harvard Univ. Press, 2009.

Chabris, Christopher, and Daniel Simons. *The Invisible Gorilla: How Our Intuitions Deceive Us*. Broadway Paperbacks, 2011.

Clavien, Christine, and Chloë FitzGerald. "The Evolution of Moral Intuitions and Their Feeling of Rightness." *The Routledge Handbook of Evolution and Philosophy*, edited by Joyce Richard, Routledge, 2018.

Cowen, Tyler. *Stubborn Attachments: A Vision for a Society of Free, Prosperous, and Responsible Individuals*. Stripe Press, 2018.

Dancy, Jonathan. *Moral Reasons*. Wiley, 1993.

Descartes, René. *Discourse on Method*. Translated by D.A. Cress, 3rd ed., Hackett Publishing Company, 1998.

Haidt, Jonathan, et al. "Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism." *Advances in Experimental Social Psychology*, vol. 47, Dec. 2012.

Hübner, Dietmar. "Three Remarks on 'Reflective Equilibrium': On the Use and Misuse of Rawls' Balancing Concept in Contemporary Ethics." *Philosophical Inquiry*, vol. 41, no. 1, 2017, pp. 11–40, doi:10.5840/philinquiry20174112.

Ichikawa, Jonathan Jenkins, and Matthias Steup. "The Analysis of Knowledge." *The Stanford Encyclopedia of Philosophy*, edited by Edward

N. Zalta, Winter 2016, Metaphysics Research Lab, Stanford University, 2016, plato.stanford.edu/archives/win2016/entries/knowledge-analysis/.

Kahneman, Daniel. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.

Liao, S. Matthew. "A Defense of Intuitions." *Philosophical Studies*, vol. 140, no. 2, 2008, pp. 247–62, doi:10.1007/s11098-007-9140-x.

Mitchell, Michael. "Science Has Next to Nothing to Say About Moral Intuitions." Aeon, 22 June 2016.

Pust, Joel. "Intuition." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2017, Metaphysics Research Lab, Stanford University, 2017, plato.stanford.edu/archives/sum2017/entries/intuition/.

Sayre-McCord, Geoff. "Moral Realism." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2017, Metaphysics Research Lab, Stanford University, 2017, plato.stanford.edu/archives/fall2017/entries/moral-realism/.

Singer, Peter. "Ethics and Intuitions." *The Journal of Ethics*, vol. 9, no. 3, Oct. 2005, pp. 331–52, doi:10.1007/s10892-005-3508-y.

—. "The Drowning Child and the Expanding Circle." *New Internationalist*, Apr. 1997.

Sinnott-Armstrong, Walter. "Consequentialism." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2015, Metaphysics Research Lab, Stanford University, 2015, plato.stanford.edu/archives/win2015/entries/consequentialism/.

Stich, Stephen. "Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity." *Synthese*, vol. 74, no. 3, 1988, pp. 391–413.

van Inwagen, Peter. "Materialism and the Psychological-Continuity Account of Personal Identity." *Philosophical Perspectives*, vol. 11, 1997, pp. 305–19.

Biography

Morgan Buzz Lawless was born on March 6, 1997 in White Plains, New York to David Lawless and Lisa Furgatch. He lived with his parents and younger sister, Samantha, in Mamaroneck, New York while attending Scarsdale Public Schools. He enrolled in the University of Texas at Austin in 2015 with majors in Plan II and computer science. In college, he studied great books in the Jefferson Scholars Program, founded Longhorns for Voting Reform, and brought ranked choice voting to student government elections. He graduated in 2019 and plans to work as a software engineer with an eye toward public policy.